April 28 2022

# Sign Speak

Lokeshwar Jha
Pranav Varma
Zhiping Liu
Afia Khan

# Table of content


01 Product Concept/ Overview

02 Product AI Canvas

03 Team

04 Value of Data Calculation for users

05 Data Flywheel

06 Model Type Selection

07 Model Metrics

08 Inferences

09 Model Results

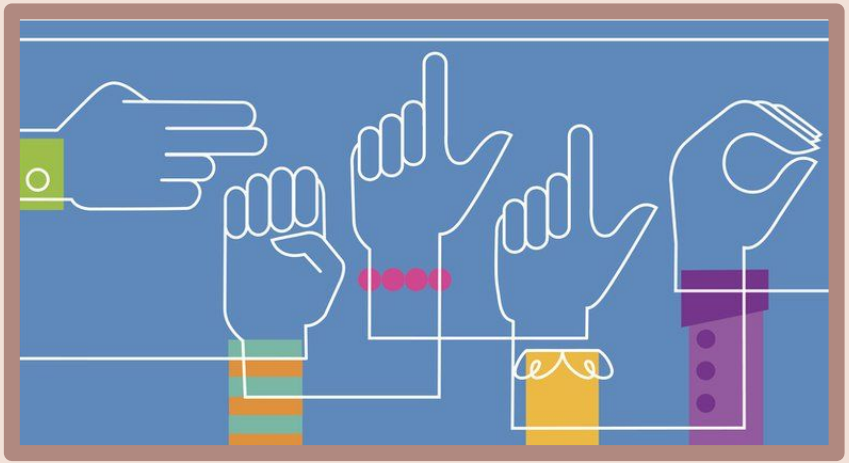
10 MVP Development

11 The Proposed Architecture

12 End-User Application

13 Governance & Monitoring

# Product Concept/Overview



To build an interactive platform through which a mute  person can communicate using sign language or text/visuals.

We envision this product to be utilized in customer-care support for specially abled and kiosks in shopping malls.

# Product AI Canvas

## AI Canvas

### Opportunity
Better user experience for customers with hearing or speaking aid

### Solution
To read user's gestures through their webcam, process the message and interact with them through text.

### Users
specially abled customers such as deaf and mute individuals

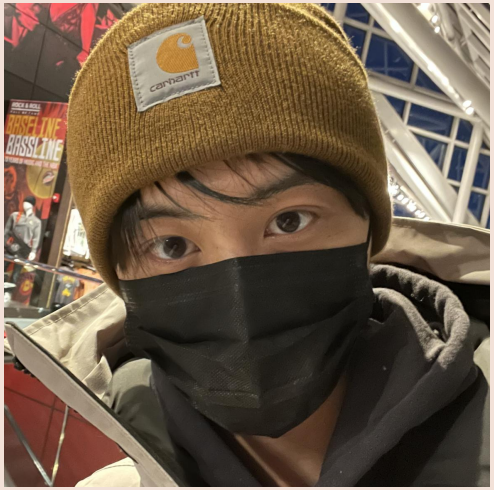
### Data
MNSIT (A-Z signs)
ASL, (0-9, A-Z signs)
WLASL (Video data)

### Strategy
Understanding the sign language and replying in text

### Success Criteria
Correct response to user

# Team

Lokeshwar




Zhiping

Pranav

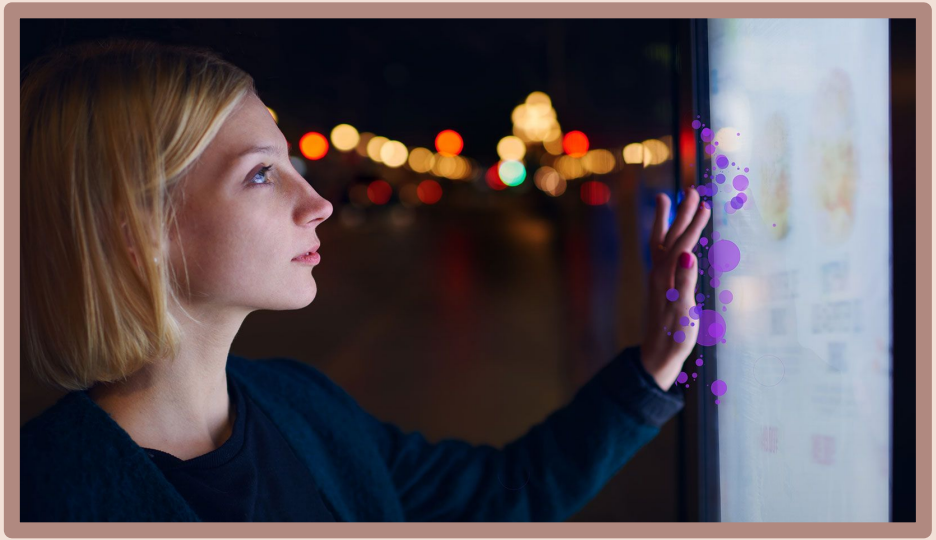
Afia

# Value of Data Calculation for User

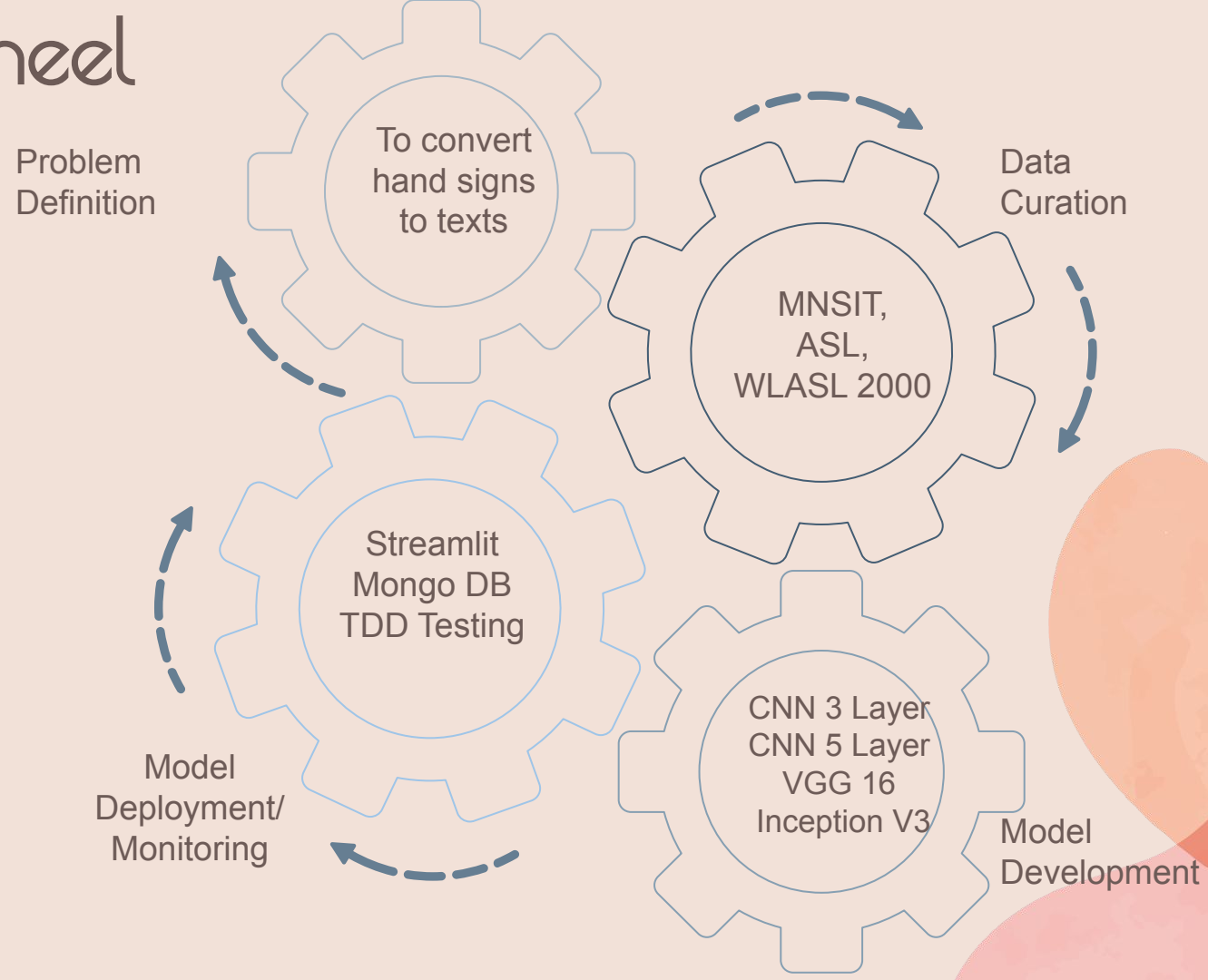
## Improving Customer Experience

Texting is one way to interact, it lacks the personalized touch, especially for companies who value customer satisfaction at their core.

This product addresses the needs of the specially abled community to make them feel more welcomed and engaged.



# Data Flywheel



# Model Type Selection

Logistic Regression

Baseline Model

CNN 3-Layers,
CNN 5-Layers

CNNs work well with image classification tasks

Transfer learning with VGG-16

Updated the last layer for our use-case while preserving the pretrained weights

Transfer learning with InceptionV3

[ Inception blocks after convolving with multiple filters seems to give us the best results by far ]

# Model Metrics

**Accuracy** Model Test Accuracy

**Precision** Precision on all the symbols, letters & digits

**Recall** Recall on all the symbols, letters & digits

## Testing in real-time scenarios

[ We tried to test the model performances on real-time image captures through streamlit app ]

# Model Results

| Models | Dataset | Accuracy |
|---|---|---|
| Logistic Regression | MNIST Dataset (28*28)<br>Alphabet Images (200*200) | Train: 77.5% I Test: 74.37%<br>Train: 72% I Test: 68.4% |
| CNN 3 Layers | MNIST Dataset (28*28)<br>Alphabet Images (200*200) | Train: 95.1% I Test: 93.2%<br>Train: 98.1% I Test: 97.3% |
| CNN 5 Layers | MNIST Dataset (28*28)<br>Alphabet Images (200*200) | Train: 96.2% I Test: 95.3%<br>Train: 98.3% I Test: 97.4% |
| VGG-16 | MNIST Dataset (28*28)<br>Alphabet Images (200*200) | Train: 99.8% I Test: 99.01%<br>Train: 97.2% I Test: 98.5% |
| InceptionV3 | MNIST Dataset (28*28)<br>Alphabet Images (200*200) | Train: 99.9% I Test: 99.7%<br>Train: 98.6% I Test: 98.9% |

# Inferences

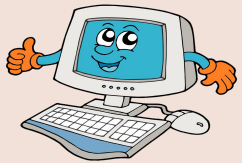
- Two signs, "M" and "S" are being confused by the model.
- Higher resolution images work well
- Pretrained models seem to work well on the datasets we trained on
- For real-time sign to text conversion, model performance worsens from the experiments
- Various data augmentation would help us in the real-time prediction
- Fingerspelling of larger words and sentences is not very suitable with the current model.

# MVP development

StreamLIT UI App

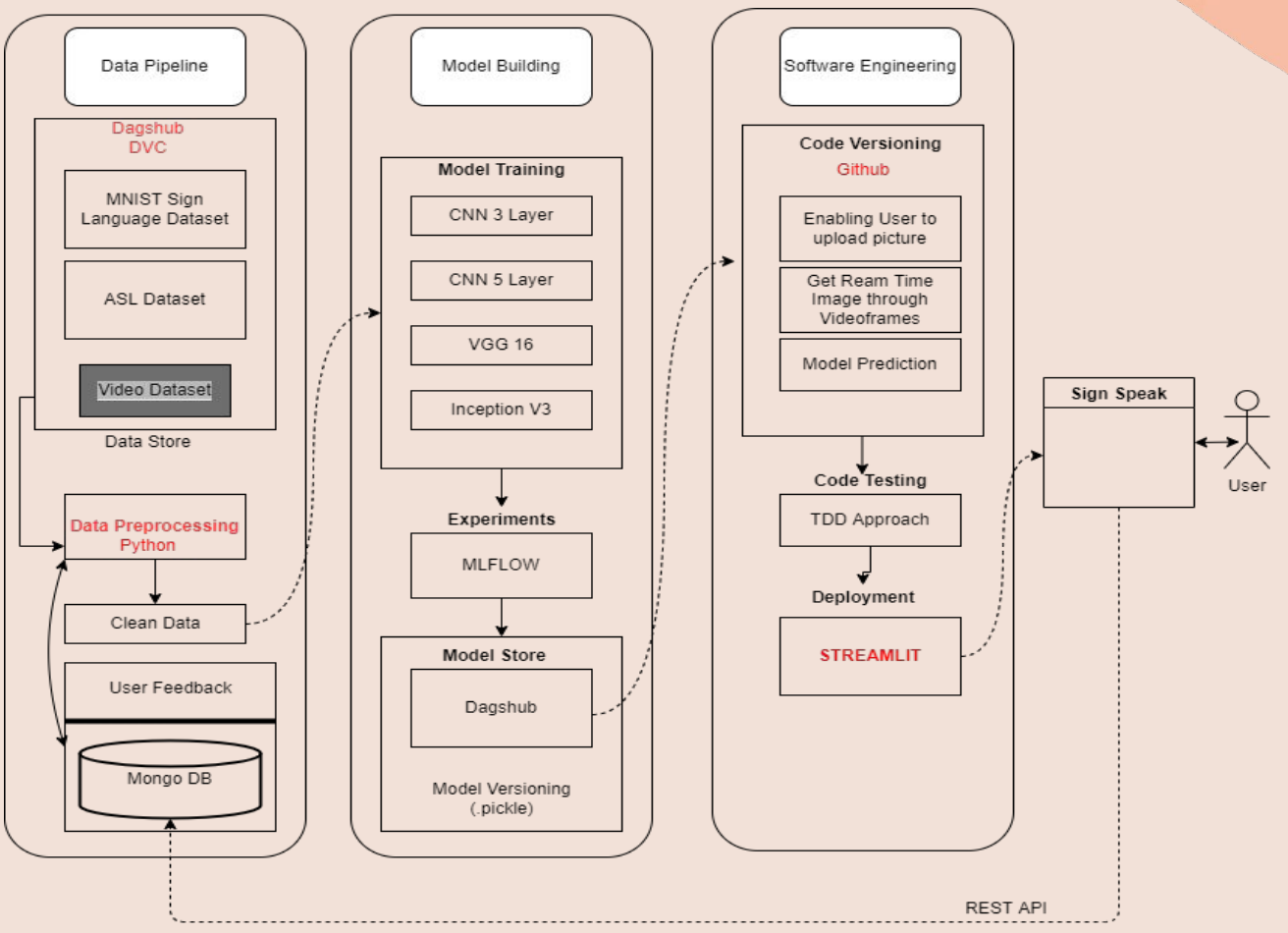
User Inputs Image of Sign

App responds with the meaning/interpretation

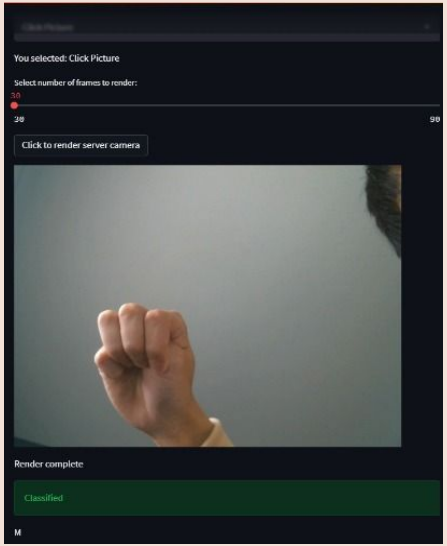


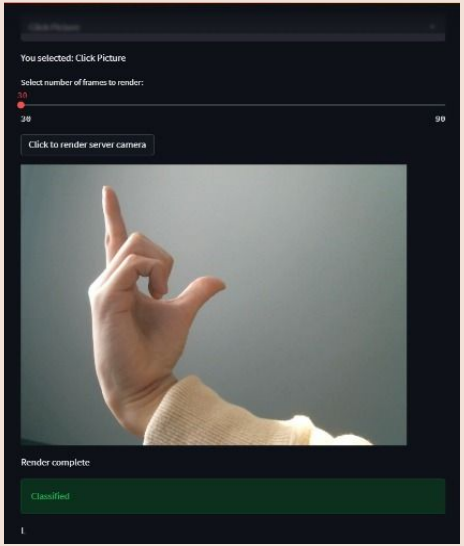
[ Product envisioned to be used by customer service orgs/kiosks to better service everyone ]

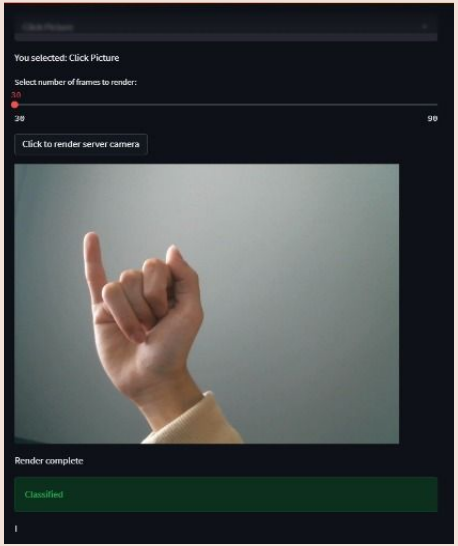
# The Proposed Architecture

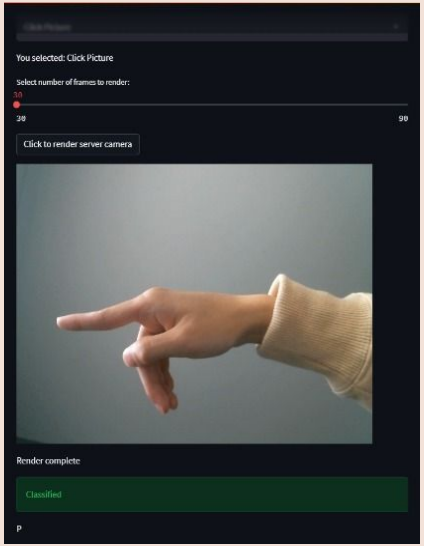


# End-user Application




M L I P

# Governance & Monitoring

- The dataset that we used is heavily biased on whites.
- Model doesn't seem to work well in other scenarios.
- Planning to scrape balanced data to further improve the model
- Sign languages are often region based. In order to make it more robust, can train for other languages
- Monitoring model and data drift through the user feedback in order to improve model performance.

# Project Repo

- Dagshub: https://dagshub.com/loki2124/SignSpeak
- Streamlit: localhost [for now]

# THANK YOU!